《心理学报》论文自检报告

请作者填写以下内容，粘贴在稿件的首页。

1. 请以“研究亮点”的形式列出最多三条本研究的创新性贡献，总共不超过 200 字。

《心理学报》的目标是发表“既科学优秀，又具有广泛兴趣和意义” (be both scientifically excellent and of particularly broad interest and significance)的心理学前沿研究。如果您的研究只有小修小补的贡献，没有尝试开创新的研究领域(new areas of inquiry)或提出独到见解和创新视角(unique and innovative perspectives)，特别纯粹只是研究没有明确心理学问题的算法或技术的工作，这类研究被本刊接受的机会小，建议另投他刊。

答：①元记忆监测的线索整合是对 Koriat (1997)提出线索利用模型的延伸，不局限于线索类型的划分，而是考察线索之间的相互作用以及个体与线索之间实际利用关系。②元记忆监测准确性对个体的记忆具有重要意义，而其神经机制确少有研究。③本研究考察情绪和吸引力两线索的整合对学习判断准确性的影响。发现线索整合存在个体差异，并且会促进学习判断准确性。ERPs 结果表明体现在编码阶段晚期能够更好地进行认知评估，在学习判断阶段能更好地进行信息保持，从而更准确地监测自身面孔记忆的认知过程。

2. 作者已经投稿或发表的文章中是否采用了与本研究相同的数据？如果是，请把文章附上审查。(我们不赞成作者用同一数据发表多篇变量相同的文章，也不赞成将一系列的相关研究拆成多个研究来发表的做法。)

答：否。

3. 管理、临床、人格和社会等领域仅有自我报告(问卷法)的非实验非干预研究，需要检查数据是否存在共同方法偏差(common method bias)。为控制或证明这种偏差不会影响研究结论的效度，你使用了什么方法？采取了哪些措施？(共同方法偏差的有关文献可参见：http://journal.psych.ac.cn/xlkxjz/CN/abstract/abstract894.shtml)基于横断数据，仅有自我报告，仅仅在方便样本中施测，这样的研究数据易取得，但通常创新性价值不大，被本刊接受的机会小。

答：本研究为实验研究，不存在上述问题。

4. 是否报告并分析了效果量(effect sizes；如：*t* 检验：Cohen's *d*；方差分析：η^2 或η_p^2；标准化回归系数)？(很多研究只是机械地报告了效果量，但没有做必要的分析或说明，如效果量是大中小？有什么理论意义或应用意义？)。(在 google 中搜索“effect size calculator”，可搜到许多计算方便的 APP。效应量的有关解释，中文可参考：http://journal.psych.ac.cn/xlkxjz/CN/abstract/abstract1150.shtml；英文可参看：http://www.uccs.edu/lbecker/effect-size.html

是否报告统计分析的 95% CI？(如，差异的 95% CI；相关/回归系数的 95% CI)置信区间的有关计算和绘图可参考 https://thenewstatistics.com/itns/esci/)

答：是。

5. 请写出计划的样本量，实际的样本量。如果二者有差别，请写出理由。以往心理学研究中普遍存在样本量不足导致的低统计功效(power)问题，我们建议在论文的方法部分解释您计算及认定样本量的依据。应该以有一定依据的效果量(effect size)、期望的功效来确定样本量，并报告计算用软件或程序。样本量计划的理由和做法可参考 https://osf.io/5awp4/

答：本研究同时考察了记忆和元记忆监测，在决定被试数量时，同时采用两种计算被试量的方法以做到相互补充。利用 G*power 参考同样考察线索对记忆和元记忆监测影响的前人研究中的效应量(Undorf & Bröder, 2020; Undorf et al., 2018)，以及另一种方法—“Power ANalysis for GEneral ANOVA”工具确定被试量，在本研究中，正式实验 1 与实验 2 中，分别选取了 41 名和 45 名被试，最终分别保留 40 和 42 个有效数据。

6. 假设检验中，如果是零假设显著性检验(NHST)，需报告精确 *p* 值而不是 *p* 的区间(小于 0.001 的报告区间，其他报告精确 *p* 值)。你的论文是否符合该项要求？如果是贝叶斯因素，是否已报告其对先验分布假定的敏感性？

答：是。

7. 为保证论文中数据报告的完备性，统计分析中如果剔除了部分数据，是否在文中报告？原因是什么？包含这部分数据的统计结果如何变化？统计分析中是如何处理缺失数据的？使用量表时是否删除了其中的个别题目？原因是什么？如果包含这部分题目，统计结果会如何变化？是否有测量的项目或者变量没有报告？原因是什么？请写出在论文中的位置。

答：是。行为实验被试剔除原因：有部分被试因个人原因中途退出；ERPs 实验被试剔除原因：在剔除脑电电压绝对值超过 ± 100 μV 的试次的时候，超过 35% 的试次数被剔除的被试不符合条件。均已在文中报告。统计结果无显著变化。

8. 研究用到的未经过同行评议和审查的实验材料、量表或问卷，是否附在文件的末尾以供审查？如果没有，请写出理由。如果该文发表，您是否愿意公开这些材料与其他研究者共享？

答：没有用到该类材料。

9. 本刊要求作者提供原始数据，请在以下 3 种里选择一种打√：

a)投稿后将数据发至编辑部邮箱 (√)

b)数据可以从如下链接中获得 ________________________ ( )

c)原始数据和程序已在心理科学数据银行(https://psych.scidb.cn/)上分享 ( )

d)如不能提供，请说明理由或提供有关证明。 ( )

10. 您的研究是否是临床干预或实验室实验？ 是(√) 否( )

如果是，请提供预注册登记号 ____________________。

如果没有，请说明原因____________________________________。

注：临床干预或实验室实验，建议在收集数据前预注册(pre-register)。也鼓励其他实验研究预注册。预注册要求写出所有的研究假设及其支持，以及实验/干预的详细过程和步骤。本期刊的预注册网站是 https://os.psych.ac.cn/preregister (使用说明书见本刊网站“下载中心”)或 https://osf.io/ 或 https://aspredicted.org/。如果您的研究有预注册，会显著增加被录用的机会。预注册的重要性可参考 https://osf.io/5awp4/

答：否，本实验开展是 3 年前，《心理学报》尚未有预注册制度。

11. 您的研究如果用到了人类或动物被试，是否得到所在单位伦理委员会的批准？如果是，请把扫描版发至编辑部邮箱。如果否，请说明理由。

答：是。

12. 是否依据编辑部网站发布的“英文摘要写作注意事项”撰写 400~500 个单词的英文大摘要？英文题目和摘要是否已请英语好的专业人士把关或者已送专业 SCI/SSCI 论文编辑公司修改润色？

答：已撰写，是。

13. 如果第一作者是学生，请导师单独给编辑部(xuebao@psych.ac.cn)发邮件，说明

已阅读本文并认真把关。是否已提醒导师给编辑部发邮件？(编辑部收到导师邮件后才会考虑进入稿件处理流程)

答：是。

14. 请到编辑部网站首页右侧“下载中心”下载并填写“稿件不涉密证明”，加盖通讯作者单位的保密办公章，把扫描件发至编辑部邮箱(xuebao@psych.ac.cn)。如没有保密办公章，请加盖通讯作者的单位公章。是否已发邮件？

答：否，投稿后会发至编辑部邮箱。

情绪和吸引力线索整合促进学习判断准确性：来自行为和脑电的证据

摘　要　线索利用模型指出个体利用不同线索进行元记忆监测，而此过程会发生线索整合。以往研究忽视面孔记忆中线索整合在元记忆监测中的作用。本研究借助影响面孔元记忆监测的情绪和吸引力两线索，探究线索整合对学习判断准确性的促进作用及其电生理机制。研究结果发现：(1)超过半数的被试整合情绪和吸引力两线索进行学习判断评分。(2)线索整合能够促进学习判断准确性。(3)线索整合组相比于非线索整合组在编码阶段顶区出现的晚期正波(LPP)与学习判断阶段额区出现的晚期负波(NSW)的波幅更大，且均与学习判断准确性呈显著相关性。研究结果表明，情绪和吸引力线索整合促进学习判断准确性，其内在机制体现在编码阶段晚期更好地进行认知评估，在学习判断阶段更好地进行信息保持，从而更准确地监测自身面孔记忆的认知过程。这一结果进一步补充线索利用模型中的个体差异，并且揭示线索整合促进学习判断准确性的神经机制。

关键词　学习判断，线索整合，元记忆

1 引言

学习判断(judgments of learning, JOLs)是指学习者对已经学过的知识在未来测试中能否成功提取的主观性预测，是元记忆监测最常用的指标之一(McDonough et al., 2021)。学习判断与记忆成绩的关系被称为学习判断的准确性(Hughes & Thomas, 2022)。准确的学习判断可以促进学习，进而使得个体能够有效调节和控制自己的心理过程(Loprinzi et al., 2022)。

研究表明，情绪和吸引力是影响面孔记忆学习判断的重要线索(Flurry, 2021; Hourihan & Bursey, 2017)。带有情绪性的面孔比中性面孔的学习判断值更高。Michaela 等人(2020)要求被试记忆带有消极和中性情绪的图片后做出学习判断。结果发现与中性图片相比，被试对消极的图片给予更高的学

习判断值。Hourihan 和 Bursey(2017)采用积极和中性的情绪图片作为学习材料，也发现个体对积极图片给予的学习判断值高于中性图片。除面孔的情绪信息外，吸引力也是面孔学习判断的重要线索(Koriat, 1997)。研究者操纵了面孔的吸引力水平，发现与低吸引力强度相比，个体对高吸引力强度和中吸引力强度面孔的学习判断准确性更高(Yuchi, & Li, 2014)。

但是，上述研究仅单独考察了面孔的情绪或吸引力线索对学习判断的影响。在日常生活中，个体对单一面孔的识别和记忆过程往往同时受到情绪和吸引力的共同影响。此外，当面孔具有情绪信息时，其吸引力水平往往也会发生改变。根据吸引力加强反应理论(reinforcement-affect theory of attraction)，微笑的面孔使人看起来更加美好，即愉快的面孔表情会增加面孔吸引力；反之，会降低面孔吸引力(Byrne & Clore, 1970)。因此，情绪和吸引力线索在面孔元记忆监测过程中可能会相互影响。这种多个线索影响元记忆监测的过程便涉及线索整合。

线索整合(cue integration)是指在非单一线索情景下，个体基于两个及以上的有效线索做出信心评级的过程(Bröder & Undorf, 2019; Undorf et al., 2018; Undorf & Bröder, 2021)。线索整合的概念是对Koriat(1997)提出的线索利用模型的延伸，将影响元记忆监测的线索不局限于类型的划分，而是重点考察线索之间的相互作用及个体在利用不同线索时的差异，其实质是个体与影响元记忆监测线索之间的交互作用。判断被试是否发生线索整合，可以通过逐一评估单个线索对某被试做学习判断是否会产生影响来进行。判断指标有两种，一是均值差异(mean difference)，计算方法是将某被试在某线索的不同水平下的学习判断值进行均值差异检验，若差异显著，则说明该被试做学习判断时基于该线索；二是计算效应量，也叫线索效应(cue effect)，计算方法是在计算均值差异的过程中，得出差异检验的效应量，只要效应量达到所规定的小效应量的标准时(并不要求均值差异显著)，则说明该被试的学习判断基于该线索。若此被试基于两种或两种以上线索做出学习判断，则说明该被试发生了线索整合。但是，效应量的实质是均值差异的大小，只有差异显著，效应量才是更有意义的(Cohen, 1988)。因此，在进行线索整合的统计时，若采用效应量的方法，则应保证均值差异显著和线索效应达到小效应量的标准，才能说明个体做学习判断时基于该线索。然而，已有研究运用线索整合的效应量指标时，并未考虑均值差异显著的前提(Undorf & Bröder, 2021)。因此，在本研究中采用更为严格的效应量指标，既保证均值差异显著，又保证效应量达到小效应量的标准，以此考察学习判断过程中的线索整合。

实际上，在多线索情境下，被试是否在学习判断中进行线索整合存在差异。以往研究根据上述均值差异或线索效应的方法区分个体是否进行线索整合，进而统计线索整合的人数所占比例(Bröder & Undorf, 2019; Undorf & Bröder, 2021)。例如，词对学习研究发现，54% 的被试基于字体大小和出

现频率线索的整合进行学习判断，但 46% 的被试并未进行线索整合，其中有的仅基于一个线索，还有的被试在做学习判断时并未基于这两个线索中的任何一个(Undorf et al., 2018)；同样地，另有研究发现大部分被试会整合词汇的具体性和情绪线索做出学习判断，研究者进一步比较线索整合组与非线索整合组的记忆成绩，发现前者更高，但并未对两组的学习判断准确性进行比较(Bröder & Undorf, 2019)；在场景图片材料的研究中，也发现了类似的结果。研究者在实验中操纵了 2 个和场景图片相关的线索，发现 46.3% 的被试整合两种线索做出学习判断，而剩余被试并未发生线索整合(Undorf & Bröder, 2021)。综上，在不同材料记忆(文字材料和场景图片)中，学习判断阶段发生线索整合均存在个体差异。因此，本研究考察面孔情绪和吸引力线索整合对学习判断准确性的影响，需要首先依据上述研究中的数据分析方法，区分出线索整合组和非线索整合组。

已有关于学习判断线索整合的研究仅从行为层面进行探究，尚未揭示线索整合影响学习判断准确性的生理机制。ERPs 研究发现，额区与学习判断过程紧密相关。词汇列表范式发现学习判断与额区至中央区出现的晚期正波相关(Skavhaug et al., 2010, 2013)。Müller 等(2016)发现类似的结果，被试在做学习判断时，在额区出现和任务相关的正慢波(positive slow wave, PSW)。因此，本研究在进行 ERPs 数据分析时，学习判断阶段脑电成分定位在额区。另外，先前研究都仅探究编码或学习判断单个阶段的脑电成分是否与学习判断存在显著相关性，而学习判断的准确性会涉及到多个阶段的神经活动。并且询问被试发现，在学习判断开始前的编码晚期阶段与整个学习阶段，被试都会进行学习判断的过程。因此，探究多个阶段与学习判断准确性相关的脑电成分能够更深入地理解被试做出信心评级的时程特征。

此外，面孔记忆相关的 ERPs 研究发现，枕区的 P1 成分以及顶区的晚期正波(later positive potential, LPP)都与面孔加工有关，前额区的晚期负走向慢波（negative-going slow wave, NSW）与面孔信息的保持有关。具体而言，P1 发生在早期阶段，主要加工低水平的面孔物理特征(Zhang et al., 2013; Zhang et al., 2014; Zhang et al., 2019)。LPP 发生在顶区，一般出现在面孔编码的晚期阶段，与情绪和吸引力有关，并能够反映高水平的注意资源分配和认知评估，也与情绪调节、标记情绪刺激进行编码和巩固有关(Cauwenberge et al., 2019; Fields, 2023)。额区的 NSW 成分出现在记忆编码后的保持阶段，和信息保持与复述相关(Reinhart et al., 2012; Wang et al., 2016)。面孔记忆相关研究中也发现前额区的负走向慢波(negative-going slow wave, NSW)和信息保持相关(Zhang et al., 2019)。以往研究从多个阶段探究面孔记忆相关的 ERPs 成分，学习判断监测记忆的过程也会涉及到多个阶段。因此，加入学习判断后，被试在各个阶段与记忆相关的 ERPs 成分可能会与学习判断及其准确性存在一定的关联。

本研究旨在探究情绪和吸引力线索整合对学习判断准确性的影响及其多阶段的时程特征。预实验目的是初步统计在学习判断阶段对情绪和吸引力线索进行整合的人数，以确保两线索操纵的有效性，进而保障后续实验的线索整合与未整合组别的人数均衡。实验 1 根据预实验结果增大吸引力线索各水平的强度梯度，计算每个被试的两个线索的效应量，据此将被试分为线索整合组和非线索整合组，进一步探究线索整合对记忆及学习判断准确性的影响。实验 2 加入 ERPs 技术，探究线索整合促进学习判断准确性的原因。学习判断准确性是记忆成绩与学习判断评分之间的相关，故学习判断准确性的提高可能会涉及编码阶段、学习判断阶段和提取阶段。由于个体提前知晓每个编码项目后都会进行学习判断。因此，个体可能在编码阶段的前期对项目进行加工记忆，而在编码阶段的晚期就已经开始进行学习判断的认知评估过程，一直持续到对信息进行复述保持的学习判断阶段，直到学习判断反应按键结束。所以，本研究采用面孔图片探究学习判断准确性的时程特征，预期会在编码阶段的晚期出现和高水平认知资源分配相关的 LPP 成分，以及在学习判断阶段出现和信息保持复述相关的 NSW 成分，这两个成分预示着学习判断准确性。而 P1 主要涉及对视觉的低水平早期加工，较少涉及更深程度更高水平的认知加工，主要和刺激的物理特征加工有关。因此，预期 P1 成分与学习判断准确性不存在相关。

2 预实验：情绪和吸引力在学习判断中线索整合水平的探究

本研究旨在探究线索整合对面孔记忆学习判断准确性的影响，选取情绪和吸引力两线索进行研究。更好地比较两线索效应的前提是保证两线索操纵方式的一致性，即选择统一的线索强度这一变量：情绪线索的操纵方式为情绪强度；吸引力线索的操纵方式为吸引力强度。预实验的目的是探究对情绪和吸引力进行线索整合和未进行线索整合的人数，以确定情绪和吸引力线索各水平的操纵是否适当，以保障后续实验的线索整合与未整合组别的人数均衡。

2.1 方法

2.1.1 被试

实验前用 G * power 软件确定被试人数(Faul et al., 2007, 2009; Legrand et al., 2021)。将 effect size 定义为 0.6 (Undorf & Bröder, 2020; Undorf et al., 2018)，α err prob 定义为 0.05，Power 定义为 0.8 时，计算出需要样本量为 16 人。

在东北师范大学选取 21 名大学生(9 男)，视力及矫正视力正常，均为右利手，且不存在精神或心理疾病。实验前用统一指导语告知被试实验目的、试验流程以及注意事项，之后签署知情同意书，完成实验后为被试发放相应报酬。因 2 名被试中途退出，最终选取 19 名(8 男)有效被试。本实验方案获得东北师范大学心理学院伦理委员会批准。

2.1.2 实验材料

实验材料均选自中国人面孔表情图片库(Gong et al., 2011)。在东北师范大学选取 24 名大学生(11 男)，对 480 张面孔图片进行吸引力强度和情绪强度 0～100 评分。将所有被试的评分进行平均，然后对每张图片进行从高分到低分的排序，选取高分 30% 为高强度，低分 30% 为低强度，中间 40% 为中等强度。最终选取 50 分以上的为高情绪强度图片，40～50 分的为中情绪强度图片，40 分以下的为低情绪强度图片。45 分以上的为高吸引力强度图片，35～45 分为中吸引力强度图片，35 分以下为低吸引力强度图片。

根据被试评阅材料的结果，共分成高情绪强度高吸引力强度、高情绪强度中吸引力强度、高情绪强度低吸引力强度、中情绪强度高吸引力强度、中情绪强度中吸引力强度、中情绪强度低吸引力强度、低情绪强度高吸引力强度、低情绪强度中吸引力强度、低情绪强度低吸引力强度 9 种面孔图片。随机选取每种面孔图片各 30 张，对效价进行平衡，正负效价各 15 张，其中 20 张用于学习阶段(学习阶段共 20×9 张面孔图片)，剩余 10 张用于测验阶段的干扰图片(测验阶段共 10×9 张新图片，20×9 张旧图片)。所有面孔图片均以相同亮度呈现于黑色背景之上，并记录所有反应。

2.1.3 实验程序

该实验程序共包含学习和测试两个阶段，见图 1。学习阶段中，共包含 9 种面孔图片，每种图片 20 张。因此，总共有 180 个 trial。在每个 trial 中，首先在电脑屏幕中央呈现“+” 500 ms，紧接着在“+”的位置呈现一张面孔图片 2000 ms ，被试的任务是尽可能记住该面孔图片。图片消失后，要求被试给出该图片的学习判断，即在屏幕上显示“你有多大把握在之后的测试阶段正确再认出这张照片？0～100 进行打分”，要求被试根据自己的真实情况用键盘直接输入 0～100 的分数以评判他们的信心，分数越高，表示被试的信心越高，即被试在之后的测试阶段越有把握再认出这张照片。输入完数字之后，按“Enter”进入下一个 trial，依此类推。

学习阶段之后，被试完成三分钟的 1000 连续减 3 的分心任务，随后进行测验。在测验阶段，总共有 270 个 trail。包括在学习阶段已学习过的 180 张旧图片和 90 张未在学习阶段呈现过的新图片，所有图片以随机顺序呈现。在每个 trial 中，首先在电脑屏幕上呈现 500 ms 的注视点，紧接着呈现 4000 ms 的图片，要求被试在 4000 ms 内判断出该图片是在之前学习阶段呈现过的旧图片(R 键)，还是未呈现过的新图片(Y 键)，并用食指按相应的按键。之后要求被试对刚才的新旧判断做 R/K 判断，即在屏幕呈现“1 记得很清楚、2 记得但不太确定、3 猜测”，要求被试根据自己的真实情况做出相应的反应。如果被试认为记得很清楚这是一张新/旧图片，则按 1；如果被试认为记得但不太确定这是一张新/旧图片，则按 2；如果被试认为新/旧判断完全是靠猜测出来的，则按 3。按键

反应之后，进入下一个 trial。数据处理去掉 R/K 判断中按 3 的 trial，排除猜测概率。

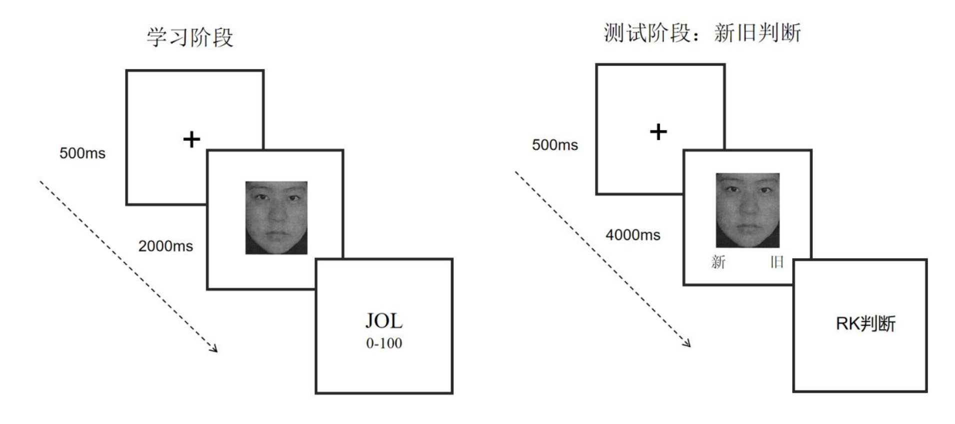


图 1　预实验流程图

2.1.4 实验设计与数据处理

数据用 SPSS 22.0 软件进行统计分析。先进行单被试分析，对每个被试分别进行情绪和吸引力线索的线索效应计算。单独求得两个线索不同水平下的学习判断评分的均值，然后对两个线索分别进行单因素三水平的重复测量方差分析：情绪线索（高强度、中强度、低强度)；吸引力线索（高强度、中强度、低强度)。若差异显著，则说明该被试具有该线索的线索效应，即在做学习判断时利用了该线索；否则，该被试并不具有此线索的线索效应，即并未利用该线索进行学习判断。对差异显著的被试，求得效应量 η_p^2 作为该线索的线索效应大小的指标。η_p^2 越大，说明该被试的线索效应越大，即在做学习判断时充分地利用了该线索。因此，部分被试可能虽然接受了实验操纵的线索，但是在做学习判断的时候可能并未利用该线索。根据每个被试是否具有情绪线索效应和吸引力线索效应，统计所有被试利用线索的情况。根据前人研究，将同时具有情绪和吸引力线索效应的定义为线索整合的被试，统计进行线索整合的被试情况。

2.2 结果分析

首先进行个体水平分析：分别对每个被试的情绪线索及吸引力线索三个强度水平的 学习判断值做方差分析，记录每个被试两线索的 η_p^2，利用 $\eta_p^2 \geq 0.02$ 的标准(Cohen, 1988)，统计被试学习判断阶段利用线索的数量，以此记录被试的线索整合情况。其中，有 84% 的被试在进行学习判断时，基于情绪和吸引力两个线索进行信心评分，有 16% 的被试仅基于情绪线索进行信心评分，未将吸引力线索作为学习判断的评分依据，见表 1。

之后，进行线索效应的补充分析。对所有被试的情绪线索的效应量和吸引力线索的效应量进行配对样本 t 检验，比较两线索的线索效应是否具有显著差异。结果发现，$t(18) = 5.53$，$p < 0.001$，Cohen'$d = 1.70$，情绪线索效应($M = 0.26 \pm 0.18$)显著大于吸引力线索效应($M = 0.04 \pm 0.03$)，说明被试对吸引力线索相比于情绪线索更不敏感。

表 1　基于不同线索做学习判断的人数(百分比)

实验	不基于两线索	仅基于情绪	仅基于吸引力	基于双线索(线索整合)
预实验	0(0%)	3(16%)	0(0%)	16(84%)
实验 1	4(10%)	9(22.5%)	5(12.5%)	22(55%)
实验 2	5(12%)	11(26%)	3(7%)	23(55%)

2.3 讨论

预实验探究面孔情绪和吸引力两个线索整合的水平。发现当情绪线索和吸引力线索同时存在时，大多数被试(84%)做学习判断的过程中更倾向于同时基于两个线索给出信心判断，即进行两个线索的整合，但是也有少部分被试(16%)仅基于情绪线索做出学习判断。这与前人研究结果一致，即在多个线索可利用的情况下，大部分被试会倾向基于多个线索进行学习判断(Undorf & Bröder, 2021)。另外，在预实验中没有发现仅基于吸引力线索做学习判断的被试。猜测可能是由于情绪线索与吸引力线索同时存在时，被试对情绪线索的敏感性更高，即被试更容易注意到面孔图片的情绪性信息，从而更大程度地根据不同情绪强度的差异做出学习判断，进而影响记忆与元记忆的信息保持和提取过程。

对线索效应的补充分析发现，学习判断阶段中吸引力线索的效应量相比于情绪线索更低，即线索效应更低，说明吸引力线索相比于情绪线索有效性更低，使得被试对吸引力线索更不敏感，进而导致了仅利用吸引力线索的人数为 0，使得线索整合组和非线索整合组的人数差异较大，难以进行线索整合与否对学习判断准确性影响的进一步分析。因此，在实验 1，增加吸引力线索三个水平之间的梯度差以便于提高吸引力线索对被试的有效性，进而保证进行线索整合的人数与未进行线索整合的人数相近，以此为基础探究被试对两个线索是否进行线索整合对记忆和学习判断准确性造成的影响。

3 实验 1：情绪和吸引力的线索整合对学习判断准确性的影响

预实验中发现被试对于吸引力线索效应相比情绪线索敏感性更低，导致线索整合组人数与非线索整合组相差较大。因此，在实验 1 中，增加吸引力线索 3 个水平之间的梯度差以便于提高吸引力线索对被试的有效性，以保证线索整合组与非线索整合组人数相当，目的是探究情绪和吸引力线索整合是否能促进记忆和学习判断准确性。

3.1 方法

3.1.1 被试

实验前用 G * power 软件进行效应量分析以此来确定被试人数。依据前人研究，将 effect size 定

义为 0.6 (Undorf & Bröder, 2020; Undorf et al., 2018)，α err prob 定义为 0.05，Power 定义为 0.8 时，计算出需要样本量为 16 人。为保证计算样本量的科学性，采用另一种方法—“Power ANalysis for GEneral ANOVA”工具(PANGEA: https://jakewestfall.shinyapps.io/pangea/; Judd et al., 2017; Westfall et al., 2015)来同时确定被试量。具体而言, 自定义设计指定了以下 4 个因子：被试(随机和交叉)， 组别(固定和交叉)，情绪强度(固定和交叉）和吸引力强度(固定和交叉)。其目的是基于 4 个指定的参数(其它参数设置为默认值)来检测各变量的主效应和交互作用。PANGEA 显示效应量 d = 0.45, 在当前设计中, 每个被试在给定实验条件下的观察次数为 20 次。该分析表明, 为确保 95%的统计效应量, 至少需要 30 名被试。

在东北师范大学选取 41 名大学生，均为右利手，视力或矫正视力正常，且不存在精神或心理疾病。实验前用统一指导语告知被试实验目的、试验流程以及注意事项，之后签署知情同意书，完成实验后为被试发放相应报酬。最终保留 40 个(18 男)有效数据。实验方案获得东北师范大学心理学院伦理委员会的批准。

3.1.2 实验材料

实验材料均选自中国人面孔表情图片库(CAFPS)，根据预实验中被试评阅材料的结果选取刺激材料。所选取的面孔图片的情绪强度梯度与实验 1 保持一致，50 分以上的为高情绪强度图片，40 ~ 50 分的为中情绪强度图片，40 分以下的为低情绪强度图片；实验 2 相比于实验 1 增大面孔图片的吸引力强度梯度，高分 25%为高吸引力强度，低分 25%为低吸引力强度，中间 25%为中吸引力强度。所以定义 60 分以上的为高吸引力强度图片，40 ~ 50 分为中吸引力强度图片，30 分以下为低吸引力强度图片。

随机选取高情绪强度高吸引力强度、高情绪强度中吸引力强度、高情绪强度低吸引力强度、中情绪强度高吸引力强度、中情绪强度中吸引力强度、中情绪强度低吸引力强度、低情绪强度高吸引力强度、低情绪强度中吸引力强度、低情绪强度低吸引力强度 9 种面孔图片，每种面孔图片各 30 张，对效价进行平衡，正负效价图片分别 15 张，其中 20 张用于学习阶段(学习阶段共 20×9 张图片)，剩余 10 张用于测验阶段的干扰图片(测验阶段共 10×9 张新图片，20×9 张旧图片)。背景呈现与预实验相同，并记录所有反应。

3.1.3 实验程序

与预实验保持一致。

3.1.4 实验设计与数据处理

实验设计采用 2(组别：线索整合组、非线索整合组) × 3(情绪线索：高强度、中强度、低强度)

× 3(吸引力线索：高强度、中强度、低强度) 的混合实验设计，其中组别为被试间变量，根据学习判断阶段的线索效应进行分组，情绪和吸引力线索为被试内变量。用 SPSS 22.0 软件进行重复测量方差分析。因变量指标包含：击中率、学习判断值、记忆成绩与学习判断的 Gamma 值、学习判断的效应量 η_p^2(线索效应)。

数据分析思路与预实验保持一致。除此之外，增加线索整合对记忆和学习判断影响的分析，即探究线索整合组与非整合组之间的差异。因此在对所有被试计算线索效应并统计完线索整合情况后，所有被试会被分为线索整合组和非线索整合组，将组别(线索整合组，非线索整合组)作为一个自变量，与情绪和吸引力线索进行混合设计方差分析，以此来探究线索整合对记忆和学习判断准确性的影响。

3.2 结果分析

3.2.1 确定线索整合组与非线索整合组人数

进行个体水平分析：记录每个被试的 η_p^2，利用$\eta_p^2 \geq 0.02$ 的标准，记录被试的线索整合情况。其中，有 22(55%)名被试在进行学习判断时，基于情绪和吸引力两个线索进行信心评分；有 9(22.5%)名被试仅基于情绪线索进行信心评分，未将吸引力线索作为学习判断的评分依据，有 5(12.5%)名被试仅基于吸引力线索进行信心评分，未将情绪线索作为学习判断的评分依据；有 4(10%)名被试未基于情绪和吸引力两线索进行信心评分。将基于两线索进行学习判断的被试划分为线索整合组，其余作为非线索整合组。见表 1。

3.2.2 线索整合对记忆和学习判断准确性的影响

基于情绪和吸引力两个线索做学习判断的被试被分到线索整合组，仅基于吸引力线索的被试为非线索整合组。将组别作为一个被试间自变量，进行 2(组别：线索整合组、非线索整合组) × 3(情绪线索：高强度、中强度、低强度) × 3(吸引力线索：高强度、中强度、低强度)的重复测量方差分析，描述统计结果见表 2。

对击中率进行重复测量方差分析，只关注与线索整合相关的结果。结果发现，组别与吸引力的交互作用显著，$F(2,76) = 4.64$，$p = 0.038$，$\eta_p^2 = 0.109$。在中等吸引力强度和低等吸引力强度的面孔图片下，线索整合组的击中率($M = 0.55 \pm 0.04$; $M = 0.46 \pm 0.04$)显著高于非线索整合组($M = 0.52 \pm 0.04$; $M = 0.46 \pm 0.04$)。在高等吸引力强度下，两组的击中率没有显著差异。组别的主效应不显著，$F(1,38) = 1.21$，$p > 0.05$，$\eta_p^2 = 0.031$。组别与情绪的交互作用不显著，$F(2,76) = 1.80$，$p > 0.05$。组别与情绪和吸引力的交互作用不显著，$F(4,152) = 1.70$，$p > 0.05$。

对学习判断准确性进行重复测量方差分析，只关注与线索整合相关的结果。结果发现，组别主效应显著，$F(1,38) = 14.84$，$p < 0.001$，$\eta_p^2 = 0.281$。线索整合组的被试学习判断准确性($M = 0.27 \pm 0.03$) 要

显著高于非线索整合组(M = 0.10 ± 0.03)。组别与吸引力的交互作用不显著，$F(2,76) = 0.89$，$p > 0.05$。组别与情绪的交互作用不显著，$F(2,76) = 0.15$，$p > 0.05$。组别与情绪和吸引力的交互作用不显著，$F(4,152) = 2.37$，$p > 0.05$。

表 2　线索整合组与非线索整合组的描述性统计结果($n_1 = 22$，$n_2 = 18$)

组别	因变量	HH	HM	HL	MH	MM	ML	LH	LM	LL
线索整合组	击中率	0.66 ± 0.19	0.63 ± 0.19	0.56 ± 0.18	0.62 ± 0.14	0.51 ± 0.20	0.54 ± 0.20	0.54 ± 0.22	0.52 ± 0.25	0.46 ± 0.29
	Gamma 值	0.19 ± 0.33	0.36 ± 0.30	0.29 ± 0.30	0.44 ± 0.33	0.23 ± 0.37	0.37 ± 0.26	0.31 ± 0.33	0.20 ± 0.45	0.04 ± 0.48
非线索整合组	击中率	0.76 ± 0.10	0.60 ± 0.13	0.48 ± 0.14	0.62 ± 0.15	0.40 ± 0.18	0.47 ± 0.17	0.49 ± 0.17	0.38 ± 0.24	0.38 ± 0.29
	Gamma 值	0.23 ± 0.40	-0.06 ± 0.44	0.16 ± 0.43	0.26 ± 0.40	0.17 ± 0.35	0.03 ± 0.27	0.13 ± 0.35	0.01 ± 0.42	-0.02 ± 0.34

注：n_1 为线索整合组的人数，n_2 为非线索整合组的人数；H 为高强度，M 为中强度，L 为低强度；第一个字母代表情绪线索，第二个字母代表吸引力线索(如：HM 为高情绪强度中吸引力强度)；表中数据表示为：$M \pm SD$

3.3 讨论

基于线索整合统计的分析结果发现，超过半数被试会同时基于情绪和吸引力线索做出学习判断，即发生线索整合。这与预实验的结果以及前人的研究保持一致(Undorf et al., 2018)。对线索整合影响记忆与学习判断准确性的分析发现，组别的击中率不显著，即线索整合组的击中率与非线索整合组并没有达到统计学上的显著差异。先前研究采用均值差异作为线索整合的指标，发现了线索整合可以提高记忆成绩，即线索整合组和非线索整合组之间击中率存在显著差异(Undorf & Bröder, 2021)。但是当采用线索效应，即效应量作为线索整合的指标时，并未发现线索整合组与非整合组在记忆成绩上的差异。这与本实验结果一致，本研究对以效应量作为线索整合指标的方法界定更为严格，使得被试在线索整合中的划分更为精确，所以也导致了击中率更不易产生差异的的结果。另外，需要关注的是，学习判断准确性在组别上主效应显著，表现为线索整合组的学习判断准确性显著高于非线索整合组。说明线索整合的被试整合了更多的学习材料的信息，以此进行了更为准确的学习判断

的信心评估，进而调整自己的学习状态和认知资源，提高了学习判断的准确性。实验 2 进一步采用 ERPs 技术，探究多个阶段与任务相关的脑电波服，以此寻求线索整合促进面孔记忆学习判断准确性的原因。

4 实验 2：情绪和吸引力的线索整合影响学习判断准确性的时程特征

实验 1 发现线索整合组的学习判断准确性显著高于非线索整合组，说明同时利用面孔情绪和吸引力两个线索相比仅利用一个线索或不利用任何线索的被试能够更有效地整合面孔的信息，以提高学习判断准确性。为揭示此现象的原因，实验 2 采用 ERPs 技术，进一步探究线索整合促进面孔图片学习判断准确性在多阶段的时程特征。为增大被试对两线索的敏感性从而容易观察脑电波幅差异，则进一步增大情绪和吸引力两线索的强度水平梯度，将 3 个水平改为 2 个水平，探究线索整合组相比于非线索整合组是否会在编码、学习判断、提取时引发更大的与情绪和吸引力线索相关的成分波幅。

4.1 方法

4.1.1 被试

实验前用 G * power 软件确定被试人数。依据前人研究，将 effect size 定义为 0.6 (Undorf & Bröder, 2020; Undorf et al., 2018)，α err prob 定义为 0.05，Power 定义为 0.8 时，计算出需要样本量为 8 人。为保证计算样本量的科学性，采用另一种方法—“Power ANalysis for GEneral ANOVA”工具(PANGEA: https://jakewestfall.shinyapps.io/pangea/; Judd et al., 2017; Westfall et al., 2015)来同时确定被试量。具体而言, 自定义设计指定了以下 4 个因子：被试(随机和交叉)， 组别(固定和交叉)，情绪强度(固定和交叉）和吸引力强度(固定和交叉)。其目的是基于 4 个指定的参数(其它参数设置为默认值)来检测各变量的主效应和交互作用(尤婷婷 等, 2023)。PANGEA 显示效应量 d = 0.45, 在当前设计中, 每个被试在给定实验条件下的观察次数为 20 次。该分析表明, 为确保 95%的统计效应量, 至少需要 30 名被试。

在东北师范大学选取 45 名大学生，对被试的要求同实验 1。最终保留 42 个(20 男)有效数据。实验方案获得东北师范大学大学心理学院伦理委员会的批准。

4.1.2 实验材料

实验材料均选自中国人面孔表情图片库(CAFPS)，根据预实验中被试评阅材料的结果选取刺激材料。选取 60 分以上的为高情绪强度图片，30 分以下的为低情绪强度图片；60 分以上的为高吸引力强度图片，30 分以下的为低吸引力强度图片。在其中随机选取高情绪强度高吸引力强度、高情绪强度低吸引力强度、低情绪强度高吸引力强度、低情绪强度低吸引力强度 4 种面孔图片，每种面

孔图片各 75 张，对效价进行平衡，正负效价图片分别为 37、38 张，其中 50 张用于学习阶段(学习阶段共 50×4 张图片)，剩余 25 张用于测验阶段的干扰图片(测验阶段共 25×4 张新图片，50×4 张旧图片)。所有背景呈现与预实验相同，并记录所有反应。

4.1.3 实验程序

为保证脑电数据的有效性，相比于实验 1，实验 2 将学习判断阶段由 0 ~ 100 按键改为两种按键：高信心、低信心。并且为保证脑电数据的成分采集，在刺激屏与反应屏后都加入 500 ms 空屏。其余与实验 1 流程保持一致。另外，正式实验开始前告知被试在实验过程中尽量避免眨眼和头动，尤其是在刺激呈现和作出反应的过程中。见图 2。

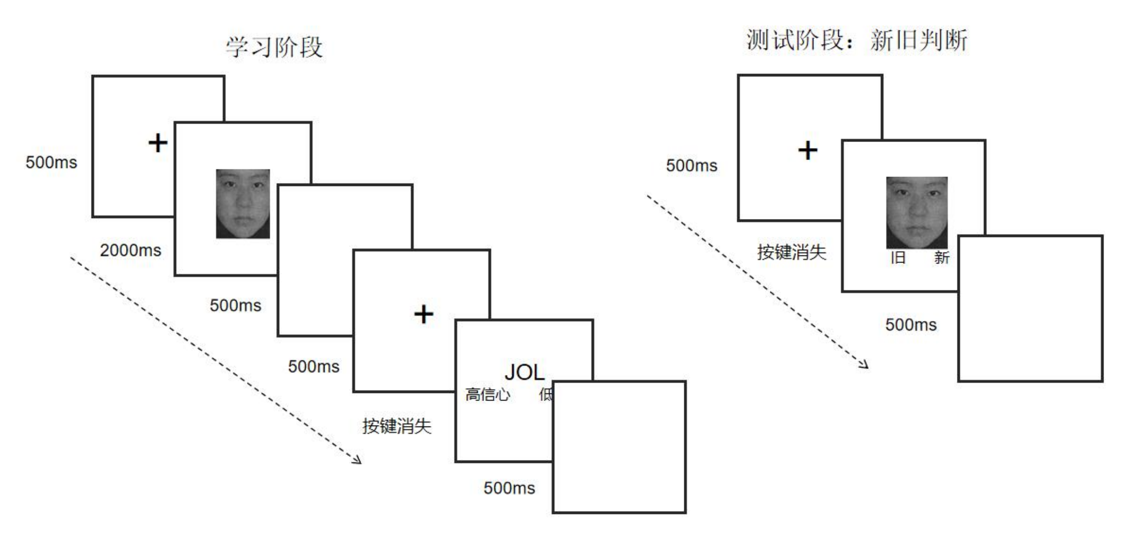


图 2　实验 2 流程图

4.1.4 EEG 记录

实验程序通过 E-prime 3.0 程序编写,通过联想 17 英寸(1024×768)的屏幕显示。要求所有被试倚靠在电脑前,眼睛距离显示器九十厘米左右。采用了 Neuroscan 公司的 SynAmps2 放大器获取在线的脑电数据,并采用了 Neuroscan Quik-Cap 64 的导电极帽,实际记录 62 个电极点的脑电数据(EEG)，电极安放位置参考国际标准 10 - 20 扩展系统。脑电实验用的左侧 M1 为在线参考电极，眼睛外侧约 1 cm 处安置水平眼电(HEOG)，左眼上下眶约 1.5 cm 处安置垂直眼电(VEOG)。所有的电极电阻在整个测试过程中都保持在 10 KΩ 以内，采样率为 1000 Hz，滤波带宽为 0.05 ~ 100Hz。

4.1.5 实验设计与数据处理

实验设计采用 2(组别：线索整合组、非线索整合组) × 2(情绪线索：高强度、低强度) × 2(吸引力线索：高强度、低强度) 的混合实验设计，其中，组别自变量为被试间变量，根据学习判断的线索效应进行分组；情绪线索和吸引力线索为被试内变量。用 SPSS 22.0 进行 ANOVA 分析。因变量指标与实验 2 一致，因变量指标包含：击中率、学习判断值、记忆成绩与学习判断的 Gamma 值、

学习判断的效应量 φ(线索效应)。

使用 Matlab 2018 软件中 EEGLAB(Delorme & Makeig, 2004)工具包进行 EEG 数据分析。首先降采样至 500 Hz，以双侧乳突的平均值(M1 + M2) / 2 进行重参考。设置 0.1 Hz 进行高通滤波，设置 40 Hz 进行低通滤波(Luck, 2014)。之后进行独立成分分析(independent component analysis, ICA)，除去眨眼等伪迹相关成分。再剔除脑电电压绝对值超过 ± 100 μV 的试次。如果发现有超过 35% 的试次数被剔除的被试，则该被试整体数据不纳入最终分析，发现 3 名被试的数据不符合条件。最终共有 42 名有效数据。

根据本实验的研究目的，同时分析编码阶段、学习判断阶段和提取阶段刺激呈现后诱发的 ERPs，刺激呈现前 200 ms 作为基线，基线至词汇呈现后 1000 ms 作为脑电分析时程(epoch)，然后根据实验条件进行分段。三个阶段均可根据实验条件划分为高情绪高吸引力、高情绪低吸引力、低情绪高吸引力、低情绪低吸引力；其中，提取阶段仅保留击中的试次。

本研究关注 NSW、LPP 以及 P1 成分，根据前人研究，LPP 选取 CP1、CPZ、CP2 三个电极点的平均值，时间窗选取 650~680 ms，NSW 成分选取 F1、FZ、F2 三个电极点的平均值，时间窗选取 760~860 ms，P1 选取 O1、OZ、O2 三个电极点的平均值，时间窗选取 90~120 ms(Fields, 2023; Wang et al., 2016; Zhang et al., 2014)。将采用 matlab 软件对于各时间窗振幅进行统计分析，进行单因素(组别：线索整合组，非线索整合组) 的独立样本 t 检验分析。

4.2 结果分析

4.2.1 线索整合对记忆和学习判断准确性的影响

进行个体水平分析：对每个被试的每个线索的学习判断进行 χ^2 检验，得到每个被试每个线索的效应量 φ，利用 $\varphi \geq 0.2$ 的标准(卢谢峰 等, 2011)，记录被试的线索整合情况。其中，有 23(55%)名被试在进行学习判断时，基于情绪和吸引力两个线索进行信心评分，有 11(26%)名被试仅基于情绪线索进行信心评分，3(7%)名被试仅基于吸引力线索进行信心评分，5(12%)名被试并未将情绪和吸引力线索作为学习判断的评分依据，结果见表 1。将基于两线索的被试划分为线索整合组，其余为非线索整合组。

基于情绪和吸引力两个线索做学习判断的被试被分到线索整合组，仅基于吸引力线索的被试和不基于两线索的被试为非线索整合组。将组别作为一个被试间自变量，进行 2(组别：线索整合组、非线索整合组) × 2(情绪线索：高强度、低强度) × 2(吸引力线索：高强度、低强度)的混合设计方差分析，描述统计结果见表 3。

对击中率进行重复测量方差分析。只关注与线索整合相关的结果。结果发现，组别的主效应不

显著，$F(1,40) = 0.17$，$p > 0.05$。组别与情绪的交互作用不显著，$F(1,40) = 0.12$，$p > 0.05$。组别与吸引力的交互作用不显著，$F(1,40) = 1.48$，$p > 0.05$。组别与情绪和吸引力的三项交互作用不显著，$F(1,40) = 1.17$，$p > 0.05$。

对学习判断准确性进行重复测量方差分析。只关注与线索整合相关的结果。结果发现，组别主效应显著，$F(1,40) = 8.01$，$p = 0.007$，$\eta_p^2 = 0.17$。线索整合组的被试学习判断准确性($M = 0.49 \pm 0.05$)要显著高于非线索整合组($M = 0.29 \pm 0.05$)。组别与情绪的交互作用不显著，$F(1,40) = 0.24$，$p > 0.05$。组别与吸引力的交互作用不显著，$F(1,40) = 0.22$，$p > 0.05$。组别与情绪和吸引力的三项交互作用不显著，$F(1,40) = 1.45$，$p > 0.05$。

表 3　线索整合组与非线索整合组的描述性统计结果($n_1 = 23$, $n_2 = 19$)

组别	因变量	高情绪 高吸引力	高情绪 低吸引力	低情绪 高吸引力	低情绪 低吸引力
线索整合组	击中率	0.74±0.18	0.63±0.23	0.63±0.23	0.52±0.24
	Gamma 值	0.34±0.39	0.32±0.31	0.32±0.37	0.17±0.41
非线索整合组	击中率	0.76±0.17	0.68±0.25	0.58±0.31	0.59±0.28
	Gamma 值	0.61±0.36	0.42±0.44	0.52±0.35	0.43±0.32

注：n_1为线索整合组的人数，n_2为非线索整合组的人数；表中数据表示为：$M \pm SD$

表 4　线索整合组和非线索整合组的 ERPs 波幅统计结果($n_1 = 23$, $n_2 = 19$)

脑电成分	线索整合组	非线索整合组
LPP	6.15 ± 4.47	3.41 ± 4.07
NSW	- 7.89 ± 3.98	-4.53 ± 5.24
P1	4.90 ± 3.95	7.28 ± 4.68

注：n_1为线索整合组的人数，n_2为非线索整合组的人数；表中数据表示为：$M \pm SD$

4.2.2 线索整合组与非线索整合组的 ERPs 波幅比较

基于学习判断准确性的组别主效应显著的行为结果，进一步探究其神经机制，分别对编码阶段、学习判断阶段以及提取阶段进行组别的独立样本 t 检验的差异分析。

首先，在编码阶段刺激呈现后 650 ~ 680 ms 的时间窗内，探究线索整合组和非线索整合组在顶区 CP1、CPZ、CP2 三个电极点所诱发的 LPP 波幅的差异。独立样本 t 检验分析发现，组别主效应显著，$t(40) = 2.06$，$p = 0.046$，Cohen’$d = 0.65$。与非线索整合组($M = 3.41 \pm 4.07$)相比，线索整合

组($M = 6.15 \pm 4.47$)诱发了更大的 LPP 波幅。波形图和地形图见图 3，各组的波幅值描述性统计结果见表 4。

在学习判断阶段刺激呈现后 760 ~ 860 ms 的时间窗内，探究线索整合组和非线索整合组在额区 F1、FZ、F2 三个电极点所诱发的 NSW 波幅的差异。独立样本 t 检验分析发现，组别主效应显著，$t(40) = 2.36$，$p = 0.023$，Cohen’$d = 0.75$。与非线索整合组($M = -4.53 \pm 5.24$)相比，线索整合组($M = -7.89 \pm 3.98$)诱发了更大的 NSW 波幅。波形图和地形图见图 4，各组的波幅值描述性统计结果见表 4。

在提取阶段刺激呈现后 90 ~ 120 ms 的时间窗内，探究线索整合组和非线索整合组在枕区 O1、OZ、O2 三个电极点所诱发的 P1 波幅的差异。独立样本 t 检验分析发现，组别主效应不显著，$t(40) = 1.86$，$p > 0.05$。波形图和地形图见图 5，各组的波幅值描述性统计结果见表 4。

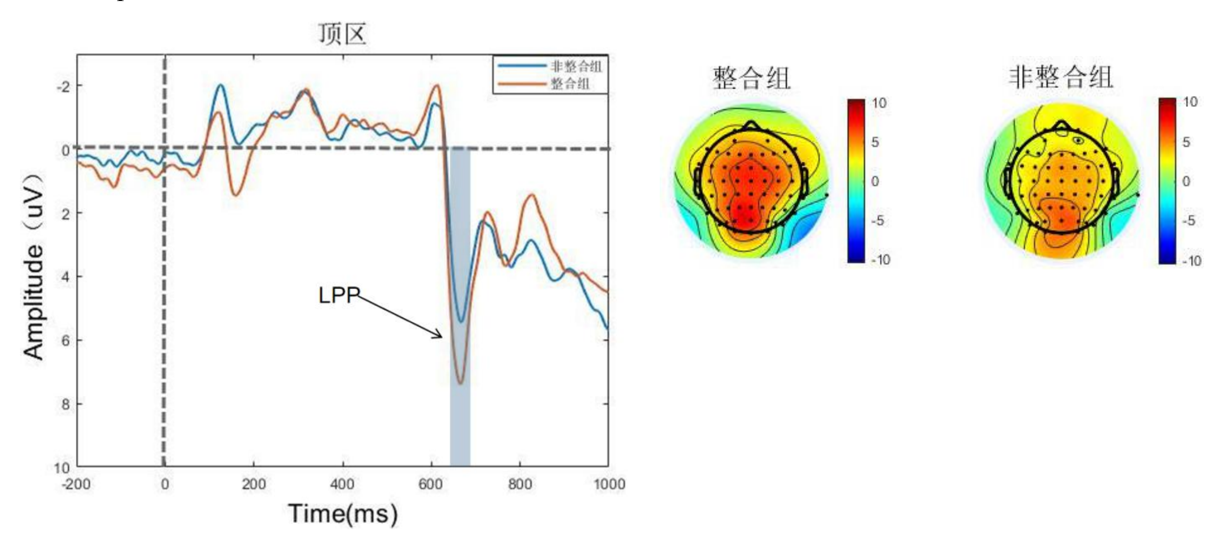


图 3 线索整合组与非线索整合组在编码阶段顶区(CP1/CPZ/CP2)的波形图和地形图

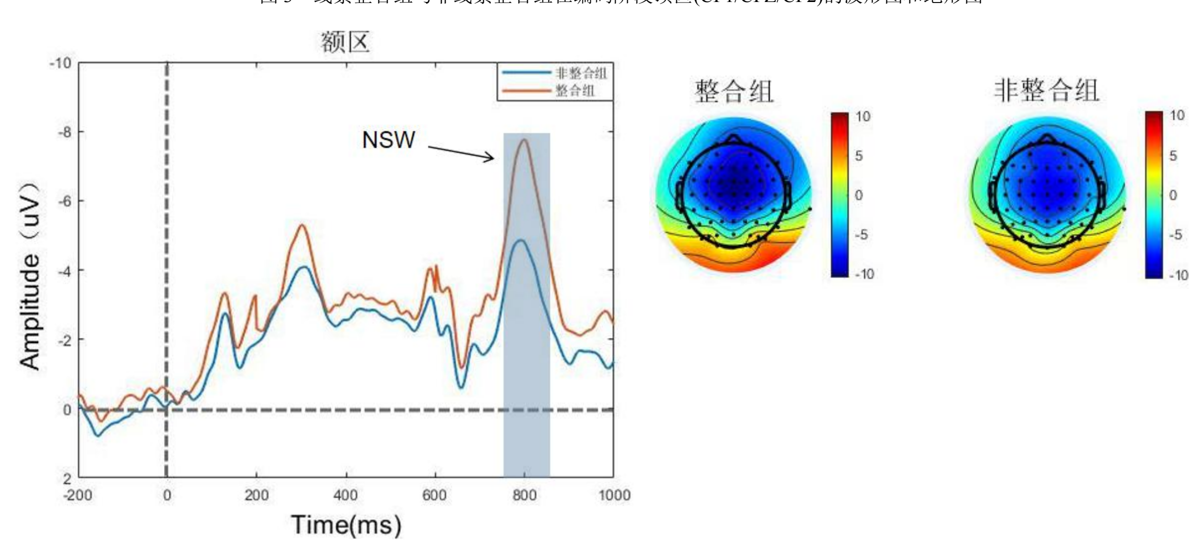


图 4 线索整合组与非线索整合组在学习判断阶段额区(F1/FZ/F2)的波形图和地形图

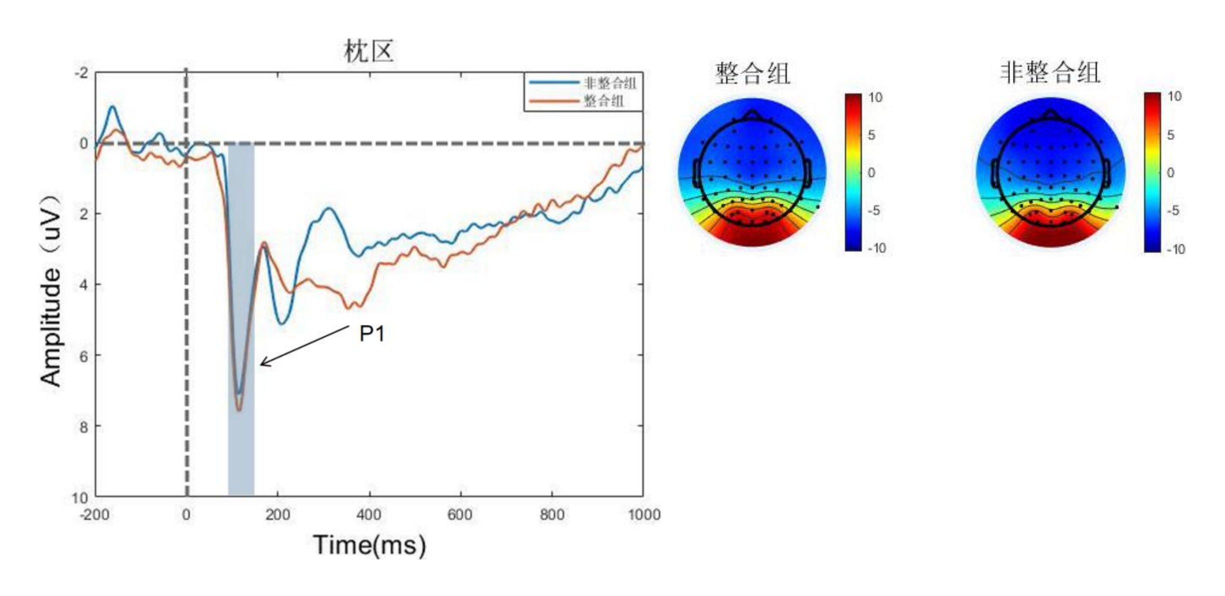


图 5　线索整合组与非线索整合组在提取阶段枕区(O1/Oz/O2)的波形图和地形图

表 5　ERPs 波幅与 Gamma 值的皮尔逊相关结果(n_1 = 23, n_2 = 19)

脑电成分	线索整合组	非线索整合组
LPP	r = 0.48, p = 0.019	r = 0.19, p = 0.43
NSW	r = 0.57, p = 0.004	r = 0.37, p = 0.12
P1	r = 0.41, p = 0.053	r = 0.03, p = 0.90

注：n_1 为线索整合组的人数，n_2 为非线索整合组的人数

4.2.3 学习判断准确性与 ERPs 波幅的相关

因为学习判断准确性这一指标会涉及多个阶段：编码阶段、学习判断阶段和提取阶段，其各阶段的脑电活动可能与此行为表现存在关联。因此，为了进一步证明线索整合和学习判断准确性的关系，本研究探讨了 ERPs 波幅与行为表现之间的联系。利用相关系数(皮尔逊 R)分析了线索整合组和非线索整合组的 LPP、NSW 以及 P1 振幅和学习判断准确性之间的相关性。如表 5 所示，结果表明，在线索整合组的被试群体中，学习判断准确性与额区的 NSW 振幅呈现显著正相关，r = 0.57，p = 0.004；与顶区的 LPP 振幅呈现显著的正相关，r = 0.48, p = 0.019；与枕区的 P1 振幅无显著相关性，r = 0.41，p > 0.05。而在非线索整合组的被试群体中，学习判断准确性与顶区的 LPP、额区的 NSW 以及枕区的 P1 振幅无显著相关性，r = 0.19, 0.37, 0.03, p = 0.43, 0.12, 0.90。该结果与 ERPs 结果一致表明，与非线索整合组相比，线索整合组的个体对面孔图片进行了更深程度的认知评估以及在学习判断阶段分配了更多的认知资源进行图片的保持和复述，诱发了更正的波幅且与行为反应正相关，而非线索整合组对图片并未根据实验中所操纵的线索进行更加细致地记忆和评估，因此学习判断准确性较低。

4.3 讨论

实验 2 的线索整合统计情况与预实验和实验 1 结果类似，超过半数被试整合了面孔情绪和吸引力两个线索，与前人研究结果一致。在线索整合影响记忆与元记忆的分析中，发现线索整合提高了学习判断准确性，但是并没有提高记忆成绩，这与实验 1 结果一致。由于实验 2 中每个线索变成 2 个水平，被试的认知负荷相比实验 1 更低，因此，不管是否整合这两个线索，被试都会有较好的成绩，所以更不易表现出两组的差异。

ERPs 结果中，在编码阶段的顶区发现与情绪认知评估相关的晚期正成分(LPP)与线索整合相关，具体表现为线索整合组的 LPP 波幅显著高于非线索整合组。说明线索整合组对于面孔图片的情绪认知评估水平相较于非线索整合组更高。在学习判断阶段的额区发现了与记忆保持相关的晚期负慢波(NSW)与线索整合相关，具体表现为线索整合的 NSW 波幅比非线索整合组更大，说明线索整合在学习判断阶段对信息进行了更深程度的复述。并且这两个成分与行为的学习判断准确性的指标存在显著相关，说明线索整合提高元记忆监测准确性的神经机制表现在额区的晚期负波和顶区的晚期正波。

5 总讨论

本研究采用经典的学习—判断—回忆范式通过三个实验，借助面孔材料探究面孔情绪和吸引力两个线索的线索整合，以及线索整合能否提高学习判断准确性及其多阶段的时程特征。预实验对情绪和吸引力线索分别操纵高、中、低三种水平，主要探究了面孔图片中被试对情绪和吸引力的线索整合水平。结果发现，有 84% 的被试进行了两线索的整合，即基于情绪和吸引力两个线索进行学习判断评定。通过情绪和吸引力线索效应的比较，发现被试对吸引力线索敏感度较低。实验 1 进一步增加了吸引力线索各水平之间的强度梯度差，以此来探究情绪和吸引力两线索的整合对记忆和学习判断准确性是否会产生影响。结果发现，有 55% 的被试进行了线索整合。另外，发现线索整合组相比于非线索整合组，学习判断准确性有明显的提升。说明当被试基于多个线索进行学习判断时，会让被试对自己记忆学习材料的认知过程有更好地监测，使得个体选择更好的学习策略进行学习，提高元记忆监测的准确性。在实验 2 中，更进一步将情绪与吸引力线索改为高、低两个水平。以此探究两线索整合对于记忆和元记忆影响的时程特征。结果发现，编码阶段在顶区出现的晚期正成分(LPP)和学习判断阶段在额区出现的晚期负成分(NSW)与行为指标的元记忆监测准确性呈现显著的相关。并且，线索整合组的 LPP 波幅相较于非线索整合组更大，说明线索整合组在编码阶段的认知评估程度比非线索整合组更深。线索整合组的 NSW 波幅相较于非线索整合组更大，说明此阶段

时线索整合组对于图片材料有更好的保持和复述，从而能够根据两线索进行更加准确的学习判断过程。

在预实验和 2 个正式实验中，都发现了大部分被试对于情绪和吸引力的两线索整合。这与以往研究结果是一致的。先前研究利用词汇材料，发现有超过一半的被试都对字词呈现频率线索和字体大小线索的整合(Undorf et al., 2018)。对场景图片的有关学习判断的线索整合研究中，实验 1 和 2 分别探究了两个内在线索和两个外在线索的整合，也都发现了超过半数的被试进行了两线索的线索整合(Undorf & Bröder, 2021)。在操纵多个线索时，每一个被操纵的线索都会成为被试做学习判断的依据。但是对于那些将所有线索都当做判断依据的被试来说，每个线索的有效性就会降低。这种结果模式与单独操纵每个线索的学习判断研究的结果有着显著的差异。根据 Kahneman(1973)编著的《注意与努力》一书中提出的认知资源有限理论，进行线索整合的被试在学习判断阶段同时关注情绪和吸引力两个线索，导致了在整体上会消耗更多的注意资源，所以对单个线索的利用程度有所降低，体现在线索效应的降低上，对每个线索的信息利用不如仅基于单线索的被试充分。

情绪和吸引力的线索整合能够促进学习判断准确性。在实验 1 和 2 中，都发现在学习判断阶段将情绪和吸引力进行线索整合的被试，他们的学习判断准确性都显著高于仅基于单个线索的被试或者不基于两线索中任意一个线索的被试。情绪和吸引力两线索之间的一些交互作用和相关性会使得被试能够整合更多的学习材料的信息，进而导致强化，以此来调整自己的学习状态和认知资源，进行更为准确的学习判断的信心评估，提高元记忆监测判断的准确性(Loprinzi et al., 2022)。另外，在实验 2 中，因为对线索的操纵划分了两个水平，所以有存在需要特征的风险，会威胁到线索对于学习判断的有效性，即当在两个容易区分的水平上操纵单个线索时，线索对学习判断的影响可能是基于被试形成怎样的预测记忆的模式会使实验更加成功的假设(Mueller & Dunlosky, 2017)。但是从学习判断准确性的结果来看，线索对于学习判断过程依然是有效的，从一定程度上排除了需求特征的影响。

学习判断准确性提高体现在编码晚期阶段更深程度的认知评估和学习判断阶段的信息保持和复述。实验 2 中，在编码阶段的顶区发现与情绪认知评估相关的晚期正成分(LPP)与线索整合相关，具体表现为线索整合组的 LPP 波幅显著高于非线索整合组。LPP 发生在顶区，与高级认知资源分配有关。例如，一项研究中指出情绪刺激相较于中性刺激诱发出更高的 LPP 幅值，以便获得更多注意资源相关(Hajcak, 2009)。一项情绪调节相关的研究发现，LPP 波幅的增大反映了注意等认知资源的增加以及与动机性刺激的深度加工。以此说明线索整合组对于面孔图片的认知资源分配及评估水平相较于非线索整合组更高(Cauwenberge et al., 2019)。在学习判断阶段的额区发现了与记忆保持相

关的晚期负慢波(NSW)与线索整合相关，具体表现为线索整合的 NSW 波幅比非线索整合组更大。发生在额区的 NSW 成分波幅越大，说明保持和复述信息的程度越深(Wang et al., 2016)。在人类和非人灵长类动物的前额叶皮质区域都发现与保持视觉信息相关的负慢波(Reinhart et al., 2012)。说明线索整合在学习判断阶段对信息进行了更深程度的复述。并且这两个成分与行为的学习判断准确性的指标存在显著正相关。与前人有关学习判断准确性的研究在额区发现的晚期负成分结果一致(Undorf et al., 2020)。

本研究首次采用更严格的线索整合的效应量指标。前人探究了对于文字材料与场景图片学习判断中的线索整合，发现了大部分被试会基于多个线索进行学习判断，并且线索整合在一定程度上提高了记忆成绩(Bröder & Undorf, 2019; Tatz & Peynircioğlu, 2019; Undorf et al, 2018)。线索整合的因变量采用效应量这一指标时，没有考虑到线索各水平均值差异显著的前提，可能会使该指标的严谨性存在不足。因此，本研究改进线索整合的效应量指标，利用面孔图片探究了情绪与吸引力线索在学习判断阶段的线索整合，发现了类似的结果，大部分被试会同时基于情绪和吸引力两个线索进行学习判断判断。并且，当人为操纵增大吸引力线索的各水平间的差异时，会提高吸引力线索的有效性，使得更多的被试在做学习判断时仅基于该线索。

对线索利用模型的支持和补充。Koriat(1997)提出将影响元记忆的监测划分为三种不同的类型，包括内在线索、外在线索和记忆线索。本研究支持线索利用模型，即个体利用线索进行元记忆监测的过程，并且不同个体利用线索的情况会有所不同，可能会受到认知资源等的影响。本研究进一步补充了线索利用模型，发现线索之间存在竞争关系，预实验中发现吸引力线索效应相比于情绪线索效应更低，导致利用吸引力线索人数更少。当在实验 1 中提高吸引力各强度水平梯度，增大吸引力线索的有效性，即发现利用吸引力线索人数增加。因此，该研究的结果在一定程度上丰富了线索利用模型，进一步印证了并不是所有线索对被试都是有效的结论，提出增加线索有效性的方法。

最后，本研究存在一定的局限性。本研究仅能证明情绪和吸引力的线索整合，且这两个线索都属于内部线索。如果线索类型更换，会不会与该实验结果的不同是不得而知的。这就导致了此次实验结果无法对两个以上线索的整合及其他线索类型进行推论。未来研究应探究面孔材料中更多数量的线索整合。另外，元记忆监测的类型单一化。本研究关注的重点是面孔学习任务中的学习判断及其准确性，因此，并未涉及其他类型的元记忆监测指标。然而，前瞻性元记忆监测的其他指标与回溯性元记忆监测同样对记忆过程非常重要(Undorf et al., 2020; Vaccaro & Fleming, 2018)。未来研究可关注线索整合领域中不同元记忆监测类型的比较。被试群体缺乏多样性。本研究选取大学生作为被试群体，他们的认知功能已经完全发育成熟，并且处于较为顶峰的时期。而学习判断作为一种高级

的认知活动，在不同的年龄阶段可能会存在着功能上的差异(Koriat et al., 2009; Liu et al., 2018)。关注其他群体在线索整合影响学习判断准确性上的行为与电生理机制的差异也是非常有必要的，为理解元记忆监测准确性的纵向发展提供一定的帮助。

6 结论

本研究通过预实验和 2 个正式实验，探讨面孔情绪和吸引力两线索的线索整合能否提高学习判断的准确性及多阶段的时程特征。得到如下研究结论：

（1）在面孔材料记忆中，超过半数被试会倾向同时基于情绪和吸引力两个线索进行学习判断，即发生情绪和吸引力的线索整合。

（2）线索整合对学习判断准确性具有促进作用，体现在编码的晚期阶段更深程度的认知评估以及学习判断阶段晚期的信息保持和复述。

参考文献

Bröder, A., & Undorf, M. (2019). Metamemory viewed through the judgment lens. *Acta Psychologica*, *19*(7), 153-165.

Byrne, D., & Clore, G. L. (1970). A reinforcement model of evaluative responses. *Personality: An international journal*.

Cauvin, S., Moulin, C. J. A., Souchay, C., Kliegel, M., & Schnitzspahn, K. M. (2019). Prospectivememory predictions in aging: Increased overconfidence in older adults. *Experimental Aging Research*, *45*(5), 436-459.

Cohen, J. (1988). The Effect Size Index: d. Statistical Power Analysis for the Behavioral Sciences.

Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, *41*(4), 1149 – 1160.

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*(2), 175 – 191.

Fields, E. C. (2023). The P300, the LPP, context updating, and memory: What is the functional significance of the emotion-related late positive potential?. *International Journal of Psychophysiology.*

Flurry, E. (2021). Facilitating multiple cue use eliminates age differences in episodic metamemory for emotional words.

Gong, X., Huang, Y. X., Wang, Y., & Luo, Y. J. (2011). Revision of the Chinese facial affective picture system. *Chinese mental health journal*.

Hajcak, G., Dunning, J. P., & Foti, D. (2009). Motivated and controlled attention to emotion: Time-course of the late positive potential. *Clinical Neurophysiology*, *120*(3), 505-510.

Hourihan, K. L., & Bursey, E. (2017). A misleading feeling of happiness: Metamemory for positive emotional and neutral pictures. *Memory(Hove, England)*, *25*, 35-43.

Hughes, G. I., & Thomas, A. K. (2022). When memory and metamemory align: How processes at encoding influence delayed judgment-of-learning accuracy. *Journal of Intelligence*, *10*(4), 101.

Kahneman, D. (1973). *Attention and effort* (Vol. 1063, pp. 218-226). Englewood Cliffs, NJ: Prentice-Hall.

Kao, Y. C., Davis, E. S., & Gabrieli, J. (2005). Neural correlates of actual and predicted memory formation. *Nature Neuroscience*, *8*(12), 1776-1783.

Koriat, A. (1997). Monitoring one's own knowledge during study: A cue-utilization approach to judgments of learning. *Journal of Experimental Psychology General*, *126*(4), 349-370.

Koriat, A., Ackerman, R., Lockl, K., & Schneider, W. (2009). The memorizing effort heuristic in judgments of learning: A developmental perspective. *Journal of Experimental Child Psychology*, *102*(3), 265-279.

Legrand, N., Engen, S. S., Correa, C. M. C., Mathiasen, N. K., Nikolova, N., Fardo, F., & Allen, M. (2021). Emotional metacognition: stimulus valence modulates cardiac arousal and metamemory. *Cognition and emotion*, *35*(4), 705-721.

Liu, Y., Su, Y., Xu, G., & Pei, M. (2018). When do you know what you know? The emergence of memory monitoring. *Journal of Experimental Child Psychology*, *166*, 34-48.

Loprinzi, P., Rigdon, B., Javadi, A. H., & Kelemen, W. L. (2022). Effects of acute exercise intensity on source episodic memory and metamemory accuracy. *Quarterly Journal of Experimental Psychology*, *75*(12), 2232-2243.

Mcdonough, M., Baillie, A. J., Clark, P. J., Gowing, L., Stapinski, L. A., & Taye, B. (2021). Understanding and managing comorbidities for people with alcohol problems: Polydrug use and dependence, co-occurring mental disorders, and physical comorbidities. *The Medical Journal of Australia*, (7), 215.

Michaela, S., A, Matthias, D. , Eva, L., & Ulrike, W. (2020). How influenceable is our metamemory for pictorial material? the impact of framing and emotionality on metamemory judgments - sciencedirect. *Cognition, 195*.

Müller, B. C., Tsalas, N. R., van Schie, H. T., Meinhardt, J., Proust, J., Sodian, B., & Paulus, M. (2016). Neural correlates of judgments of learning-An ERP study on metacognition. *Brain research*, *16*(52), 170-177.

Mueller, M. L., & Dunlosky, J. (2017). How beliefs can impact judgments of learning: Evaluating analytic processing theory

with beliefs about fluency. *Journal of Memory and Language*, *93*, 245-258.

Reinhart, R. M., Heitz, R. P., Purcell, B. A., Weigand, P. K., Schall, J. D., & Woodman, G. F. (2012). Homologous mechanisms of visuospatial working memory maintenance in macaque and human: Properties and sources. *Journal of Neuroscience*, *32*(22), 7711-7722.

Skavhaug, I. M., Wilding, E. L., & Donaldson, D. I. (2010). Judgments of learning do not reduce to memory encoding operations: Event-related potential evidence for distinct metacognitive processes. *Brain Research*, *13*(18), 87-95.

Skavhaug, I. M., Wilding, E. L., & Donaldson, D. I. (2013). Immediate judgments of learning predict subsequent recollection: Evidence from event-related potentials. *Journal of Experimental Psychology Learning Memory and Cognition*, *39*(1), 159-166.

Tatz, J. R., & Peynircioğlu, Z. F. (2019). Judgments of learning in context: Backgrounds can both reduce and produce metamemory illusions. *Memory and Cognition*, *48*, 581-595.

Undorf, M., Amaefule, C. O., & Kamp, S. M. (2020). The neurocognitive basis of metamemory: using the N400 to study the contribution of fluency to judgments of learning. *Neurobiology of Learning and Memory*, *169*, 107176.

Undorf, M., & Bröder, A. (2021). Metamemory for pictures of naturalistic scenes: Assessment of accuracy and cue utilization. *Memory and Cognition*, *49*(7), 1405-1422.

Undorf, M., Söllner, A., & Bröder, A. (2018). Simultaneous utilization of multiple cues in judgments of learning. *Memory and Cognition*, *46*(4), 507-519.

Vaccaro, A. G., & Fleming, S. M. (2018). Thinking about thinking: A coordinate-based meta-analysis of neuroimaging studies of metacognitive judgements. *Brain and neuroscience advances*, *2*, 1–14.

Wang, L., Gui, P., Li, L., Ku, Y., Bodner, M., Fan, G., & Dong, X. (2016). Neural correlates of heat-evoked pain memory in humans. *Journal of Neurophysiology*, *115*(3), 1596-1604.

You, T., Zhang, L., Qi, G., & Long, C. (2023). Opportunity (in) equity affects outcome evaluation at an early cognitive stage. *Acta Psychologica Sinica*, *55*(12), 1997.

Yuchi, Z., & Li ping. (2014). Effect of face attractiveness on face recognition and judgments of learning. *Science of Social Psychology*.

Zhang, D. D., He, W. Q., Wang, T., Luo, W. B., Zhu, X. G., Gu, R. L., Li, H., & Luo, Y. J. (2014). Three stages of emotional word processing: An ERP study with rapid serial visual presentation. *Social Cognitive and Affective Neuroscience*, *9*(12), 1897-1903.

Zhang, D. D., Luo, W. B., & Luo, Y. J. (2013). Single-trial ERP analysis reveals facial expression category in a threestage scheme. *Brain Research*, *15*(12), 78-88.

Zhang, D., Lin, Y., Liu, Y., Luo, Y., & Jiang, D. (2019). Memory encoding, retention and retrieval of disgusting and fearful faces. *Acta Psychologica Sinica*, *51*(1), 36.

Cue-integration of Emotion and Attraction Facilitates Accuracy of JOLs: the Evidence from Behavior and ERP

Abstract

Judgments of learning (JOLs) refer to learners' subjective predictions of whether they can successfully extract what they have learned on future tests. Face memory is an important foundation for acquiring information during social interactions and ensuring that social activities are carried out properly. Emotion and attractiveness are two important cues that influence JOLs of face memory. In reality, emotion and attractiveness often appear simultaneously in the same face. However, previous studies have only examined the effects of the two cues on JOLs individually, and have not deeply explored the effects of the integration of the two cues on the accuracy of JOLs and their mechanisms.

The present study first explored the proportion of the number of emotional and attractive cue integrators. Then, we increased the gradient of each level of the attractiveness cue in Experiment 1, and utilized a mixed experimental design of 2 (group: cue-integrated group, non-cue-integrated group) × 3 (emotional cue: high-intensity, medium-intensity, and low-intensity) × 3 (attractiveness cue: high-intensity, medium-intensity, and low-intensity) to explore whether the cue-integration could improve face memory and the accuracy of JOLs. In Experiment 2, in order to further improve the sensitivity of the subjects to the two cues, the mixed experimental design was changed to 2 (group: cue-integration group, non-cue-integration group) × 2 (emotional cues: high intensity, low intensity) × 2 (attraction cues: high intensity, low intensity), and the EEG was used to investigate the temporal characteristics of cue-integration in promoting the accuracy of JOLs.

Findings: (1) Subjects integrated both emotion and attraction cues for JOLs ratings(the pre-experiment). (2) Integrating emotional and attractiveness cues improved the accuracy of JOLs (Experiment 1). (3) The group that integrated cues had a higher amplitude of late positive waves (LPP) in the parietal region during the encoding stage and late negative waves (NSW) in the frontal region during the JOLs stage compared to the group that did not integrate cues. Additionally, the amplitudes of NSW and LPP in the cue-integrated group were significantly correlated with the accuracy of JOLs(Experiment 2). The study found that individuals who integrated emotion and attractiveness cues during the encoding stage were better able to allocate cognitive resources for cognitive assessment and retain information in the JOLs stage. This led to more accurate monitoring of their own face memory. The study suggests that integrating

two cues can improve cognitive performance.

This study offers a foundation for individuals to comprehend the impact of cue integration on memory and metamemory in real-world face learning scenarios. It also aids in the development of effective learning plans and strategies, as well as precise monitoring of the learning process.

Key words: judgments of learning, cue integration, metamemory